Versão preliminar
Fevereiro de 2003

As avaliações de nova geração nas sociedades contemporâneas

Simon Schwartzman

(publicado em Alberto de Melo e Souza (org), Dimensões da Avaliação Educacional, Petrópolis, Vozes, 2005, pp. 15-34)

As avaliações de nova geração nas sociedades contemporâneas

Simon Schwartzman

Educação e avaliação sempre andaram de mãos dadas. O trabalho do educador é transformar o estudante naquilo que não era, desenvolvendo competências e habilidades, inculcando valores e ampliando seus conhecimentos. O educador precisa saber sempre se o estudante está aprendendo e se desenvolvendo, e os estudantes se miram no espelho de seus mestres para saber se estão de fato progredindo, formando assim sua auto-imagem. Esta avaliação permanente ocorre em muitas atividades humanas, que estão sendo sempre sopesadas, de forma implícita ou explícita, em termos de sua qualidade, benefícios, custos, beleza ou prazer. O que é peculiar à educação – mas, novamente, não só a ela – é o desenvolvimento de sistemas complexos e sistematizados de avaliação, e o uso regular de seus resultados para diversos fins. A forma que estes sistemas assumem, e o papel que eles desempenham, dependem de muitos fatores, que variam de uma época a outra, e de uma a outra sociedade: suas tradições culturais e intelectuais, seu regime político-institucional, e os diferentes papéis que a sociedade atribui e espera da educação, em seus diversos níveis.

As funções tradicionais da educação: cimento e filtro social

A função mais tradicional da educação formal, ressaltada por sociólogos clássicos como Émile Durkheim e Talcott Parsons, é de transmitir valores e manter a coesão social, necessários sobretudo nas sociedades modernas e complexas. Durkheim, na França do início do século XX, contrastava as sociedades tradicionais, onde a coesão se daria pela similaridade das pessoas (*solidarité mécanique*) com as sociedades complexas, aonde as pessoas se complementariam pela divisão do trabalho social (*solidarité organique*). O jogo dos interesses individuais, dos contratualistas de antes e de hoje, não lhe parecia suficiente para manter a sociedade coesa. Era necessário,além disto, preservar e transmitir a cultura e dos valores, que no passado havia sido o papel da religião, e que, nas sociedades modernas deveria ser função dos sistemas educacionais. Esta idéia seria retomada décadas mais tarde nos Estados Unidos por Talcott Parsons, que atribuía às instituições educacionais a responsabilidade por uma das quatro funções

centrais de qualquer sistema social viável, a de manutenção de valores e padrões de comportamento.[1]

Este tema, do papel da educação como cimento social, ficou por muitos anos fora de moda, superado por outras questões, como as necessidades do mercado de trabalho ou de equidade e justiça de acesso. Nos últimos anos, no entanto, o tema tem sido retomado com grande interesse, a partir dos estudos e pesquisas sobre a importância do "capital social", que seria a base sem a qual as instituições políticas e os mercados não poderiam funcionar (Putnam, Leonardi, and Nanetti 1993; Dasgupta and Serageldin 2000; Fukuyama and IMF Institute 2000).

A outra função tradicional e clássica da educação é a de filtrar as pessoas para posições de prestígio e elite, conforme sua maior ou menor aderência aos cânones dominantes. O exemplo histórico mais famoso talvez tenha sido o do Mandarinato chinês, em que os *literati* se preparavam por anos, estudando literatura e caligrafia, para um exame que os levaria a uma das poucas posições de prestígio e poder na burocracia imperial. Não havia relação clara entre estes estudos e o trabalho que fariam depois. Esta pouca relação entre o conhecimento adquirido e as funções a serem desempenhadas no trabalho persiste até hoje, nos sistemas educacionais modernos. Como nos esportes ou nas artes, as pessoas não são avaliadas, somente, para saber se elas conseguem desempenhar de forma adequada suas funções, mas, também, para saber se elas são melhores ou piores que seus concorrentes.

Esta função de filtragem pode ser bastante perversa, pois dá às pessoas uma ilusão de mobilidade social que depende do descenso de outras. Daí a noção, popularizada por alguns autores, de que a educação formal não passaria de um mecanismo de consolidação e reprodução da ordem estabelecida (Bourdieu and Passeron 1970). Na realidade, o fortalecimento dos sistemas de mérito e a valorização da educação fazem parte de uma profunda revolução que substituiu as formas mais tradicionais de hierarquia e prestígio social baseados no sangue e na propriedade da terra por uma nova hierarquia baseada

---

[1] As outras funções seriam as de estabelecimento de objetivos e metas (governo), adaptação ao ambiente (economia) e integração (justiça). As referências clássicas são (Durkheim 1967);.(Durkheim and Halbwachs 1938); (Parsons and Bales 1966); (Parsons 1991)

na competência e no desempenho individual. O fato de que tenha havido compromissos e acomodações de toda ordem entre as antigas e as novas elites não diminui a importância desta transformação. Uma vez implantados os sistemas de mérito, no entanto, sua perpetuação pode fazer com que seus efeitos perversos adquiram maior relevo. Para os indivíduos, competir por um melhor lugar através da educação pode ser a melhor estratégia, principalmente para os que não são herdeiros de fortunas e filhos das antigas aristocracias. Para a sociedade como um todo, o crescimento sem limites da escolaridade formal e das exigências educacionais podem não trazer os benefícios que freqüentemente se imagina.[2]

As avaliações na educação tradicional

A implantação e manutenção destes grandes sistemas de formação, endoutrinação e filtragem das pessoas, por grandes burocracias públicas, não poderia ser feita sem mecanismos que garantissem, de alguma forma, que seus objetivos fossem atingidos. Na entrada do século XX, a maioria dos países da Europa ocidental, assim como o Japão e a Argentina, já haviam conseguido universalizar a educação básica, e avançavam a passos rápidos na implantação e universalização da educação secundária.

Em todos os países em que os governos centrais tiveram papel preponderante no financiamento e coordenação da educação básica e fundamental, os sistemas tradicionais de avaliação por consenso evoluíram para a busca de padrões nacionais que não dependessem exclusivamente das preferências e orientações subjetivas de cada professor e escola. As manifestações mais claras desta tendência foram os exames de conclusão da educação secundária, dos quais o *Baccalauréat* francês e o *Abitur* alemão, estabelecidos no século XIX, são os exemplos mais conhecidos. Em ambos os casos, as provas são acompanhadas por procedimentos extremamente complexos, de juizes e comissões de professores, que buscam fazer com que um determinado conceito dado por uma escola seja equivalente, em conteúdo, a conceitos equivalentes dados em todas as outras escolas. Estes exames são a culminação de sistemas permanentes de coordenação. acompanhamento e

---

[2] Este ponto é levantado, sobretudo para a educação superior de massas, em (Wolf 2002).Veja também, para uma perspectiva controversa, (Herrnstein and Murray 1994).

inspeções das escolas, que procuram garantir não somente que a qualidade da educação seja mantida, mas que todos os estudantes obtenham o mesmo tipo de educação. Assim, os resultados das provas finais seriam uma expressão pura das qualidades de cada indivíduo, independentemente de sua origem ou condição social, e serviriam para definir suas chances e lugar na hierarquia de oportunidades que se abrem aos de melhor desempenho, seja na universidade, seja na vida profissional.[3]

Estes sistemas foram concebidos em uma época em que poucos estudantes chegavam à escola secundária, e foram sendo flexibilizados na medida em que a educação média ia se expandindo e se diversificando. Hoje, na Alemanha, o ensino secundário acadêmico que conduz ao *Abitur* só é freqüentado por cerca de 1/4 dos estudantes, enquanto que a maioria segue cursos profissionais e técnicos de outro tipo. Na França, o antigo *Bac* acadêmico se transformou em um sistema diversificado, cada vez mais complexo, e de futuro incerto.

As novas funções da educação

Os rigorosos sistemas de controle e certificação estabelecidos pelos países europeus, propagados depois para suas colônias e áreas de influência e adotados por alguns países da Ásia, garantiram a superioridade de sua educação formal sobre a dos demais países, inclusive em relação aos Estados Unidos, país aonde o ensino sempre foi fortemente descentralizado, e o setor privado sempre teve grande presença. Além de transmitir valores e atitudes, eles transmitiam também os conhecimentos e conteúdos que faziam parte da "norma culta" dos respectivos países, a começar pelo domínio da linguagem escrita e falada. Esta vantagem da educação tradicional pode ser vista ainda hoje nas pesquisas comparadas sobre o desempenho dos estudantes em provas de conhecimento em competências em matemática: como regra geral, quanto mais tradicional e formalizado o ensino, melhor o desempenho dos alunos (OECD Programme for International Student Assessment 2001).

---

[3] Para um histórico da evolução dos sistemas públicos de avaliação, veja (The World Bank Group 2002). Para uma descrição detalhada dos diferentes sistemas nacionais de conclusão da educação secundária - sobretudo o Abitur alemão e o Baccalauréat francês, (McLauchlan de Arregui et al. 2001).

Ao longo do século XX, duas transformações profundas ocorreram em relação à educação. Primeiro, ela continuou se expandindo, chegando à universalização do acesso ao nível médio nos países mais desenvolvidos por volta da metade do século, e avançando pela extensão crescente da educação superior e da pós-graduação. Depois, educação e cultura deixaram de ser vistas, simplesmente, uma forma de socialização nos estilos, formas de vida e valores dominantes, e passaram a ser considerados como um fator de produção e riqueza. Hoje, com a vulgarização da chamada "teoria do capital humano", isto parece ser uma coisa óbvia, mas não o era no passado.

Evolução dos sistemas de educação média e superior

	Sistemas tradicionais	sistemas contemporâneos
cobertura da educação secundária	limitada	abrangente e com tendência à universalização
conteúdos da educação secundária	homogêneos e orientados para a formação humanística e científica de tipo acadêmico	muito herogêneos, com cursos de formação científica, técnica, humanística, profissional e vocacional
cobertura da educação superior	limitada: cerca de 5% da população	de massas: 30% ou mais, com tendência à universalização
conteúdos da educação superior	as profissões clássicas: medicina, direito,	altamente diferenciada: vocacional, geral, profissional, científica, educação continuada

A expansão, o aumento dos custos e a relevância econômica cada vez evidente da educação fizeram com que os antigos cânones de qualidade e os respectivos mecanismos de controle e avaliação da educação começassem ser questionados. A quebra dos cânones se deu, sobretudo, pelo rompimento da barreira que antes parecia dividir com tanta nitidez a educação acadêmica, das elites, da educação técnica e profissional. A formação buscada pelas escolas e universidades tradicionais sempre foi a das profissões clássicas, como o direito, a medicina, a teologia e o próprio magistério, cujos melhores estudantes encontrariam posições de prestígio e reconhecimento e as rendas decorrentes em função de sua competência, sem ter que competir em mercados dominados pela busca desavergonhada do lucro. Esta separação e oposição entre a formação e o treinamento técnico e profissional, para o mercado, e a educação acadêmica, para as posições de prestígio no governo e na cultura, ocorreu por toda parte, e ajudou a prolongar a auto-suficiência das instituições educacionais. Na Inglaterra, como na Alemanha, a formação técnica e

especializada foram mantidas por muito tempo fora das escolas tradicionais e das universidades, localizadas em centros de formação profissional e politécnicos, mantidos muitas vezes pelas próprias indústrias e corporações profissionais. No Brasil, a formação técnica e profissional sempre esteve, desde os anos 30, dominada pelas instituições patronais, enquanto que as Secretarias e o Ministério da Educação cuidavam (ou deveriam cuidar) da educação geral

Na medida em que esta separação entre as pretensões mais grandiosas da cultura tradicional e a realidade da vida prática da indústria e do comércio ocorria, as instituições educacionais e universitárias começaram a perder o papel central de formação e filtragem das elites que acreditavam desempenhar. Na medida em que as elites se transformavam, com a inclusão crescente de pessoas que não haviam passado pela educação convencional; na medida em que comerciantes, engenheiros, sindicalistas e militares assumiam posições de elite e de comando, trazendo seus próprios valores e perspectivas, as instituições educacionais se isolavam, e perdiam os critérios externos de lhes serviam como pontos de referência e de apoio.

Estas transformações, que se deram na Europa ao longo do século XIX, tiveram impacto profundo nos sistemas educacionais dos diversos países. Na França, a revolução e o regime napoleônico varreram com as universidades tradicionais, e entronizaram os engenheiros e as “grandes écoles” como a nova elite do Estado moderno. Na Inglaterra, as escolas e universidades tradicionais se mantiveram isoladas, servindo à aristocracia e ao governo, enquanto que a burguesia e o mercado desenvolviam suas próprias instituições e sistemas de ensino; na Alemanha, as universidades se mantiveram, por mais tempo que em qualquer outro país, como instituições ligadas ao Estado e formadoras de suas elites, até que, com a República de Weimar e o nazismo, o edifício viesse abaixo. Nos Estados Unidos, sem tradição acadêmica, a educação se desenvolveu de forma muito mais confusa e misturada, com uma combinação de valores, atitudes e conteúdos acadêmicos e pragmáticos, públicos e privados, de alta e de baixa cultura, que deixariam os europeus escandalizados. E por isto, talvez, a educação universitária americana tenha se tornado muito mais dinâmica e criativa, a longo prazo, do que na Europa.

As avaliações de nova geração

Os novos sistemas de avaliação se desenvolveram principalmente a partir dos Estados Unidos, como reação aos óbvios problemas de falta de qualidade de muitas de suas instituições educacionais, na ausência de uma corporação profissional de educadores suficientemente estruturada para tomar a si estas funções, como havia ocorrido na Europa, e com a preocupação crescente com a questão das igualdades de oportunidade, centradas sobretudo na questão racial. O marco desta nova linha de trabalhos foi o famoso "Coleman Report", de 1966 (Coleman et al. 1966), um grande estudo envolvendo 600 mil crianças em 4 mil escolas, feito por solicitação do governo norte-americano para entender o que estava ocorrendo nesta área. O relatório, baseado em análises estatísticas complexas, constatou, entre outras coisas, que as diferenças de desempenho dos estudantes não dependiam somente das características das escolas, mas, sobretudo, as condições sócio-econômicas e culturais em que viviam suas famílias, e que afetavam a população negra de forma especialmente cruel.

O relatório Coleman deu origem a muitos estudos e pesquisa sobre a questão educacional, e em 1969, por decisão do Congresso americano, foi criado o National Assessment of Educational Progress – NAEP (Avaliação Nacional do Progresso em Educação), que serviu de modelo para o Sistema de Avaliação da Educação Básica (SAEB) no Brasil. O NAEP é também conhecido como o "Boletim Escolar da Nação", e acompanha a evolução da qualidade da educação norte-americana nas diferentes regiões e Estados (National Assessment Governing Board 2002).

Além do problema da igualdade de oportunidades, os americanos se deparavam com a questão da qualidade de sua educação. Em 1981 um novo documento de grande repercussão foi publicado, "A Nation at Risk", baseado em um grande número de estudos especialmente encomendados por uma comissão de alto nível, e que culminava em uma série de recomendações para melhorar a educação do país (United States Department of Education and Education 1983)

As metodologias adotadas são muito distintas das anteriores, que eram baseadas na busca do consenso entre professores sobre conteúdos e critérios

de avaliação. Agora, as opiniões e avaliações dos especialistas continuam sendo tomadas em conta, mas são processadas e interpretadas com a ajuda de procedimentos estatísticos complexos, manejados por estatísticos, psicometristas e sociólogos que não trabalham necessariamente em instituições educacionais, e sim para órgãos de governo, institutos de pesquisa e para as novas agências de avaliação.

As funções da avaliação também são mais amplas do que no passado. Antes, as avaliações tinham por objetivo certificar se o aluno cumpria os requisitos necessários para ingressar em uma universidade, e, posteriormente, em uma das profissões clássicas. Havia a suposição de que todos os alunos que haviam passado pelos sistemas de certificação como o *Abitur* ou o *Bac* tinham a mesma qualificação, dentro de um âmbito de variação dado por suas notas. Os professores não eram avaliados, a não ser no momento de sua formatura, ou quando apresentavam sua teses para bancas ou comitês de especialistas. Hoje, estas equivalências não são mais aceitas como dadas, e precisam ser comprovadas estatisticamente. Antes, as avaliações eram sempre dos estudantes, individualmente; hoje, busca-se avaliar as instituições, o desempenho dos professores, os métodos de ensino, os programas governamentais de expansão e melhoria da educação e seu impacto; entender os condicionantes sociais dos bons e maus resultados; e identificar procedimentos que possam melhorar os resultados.

A evolução das avaliações de conclusão do ensino médio

Tradicionalmente, estas são avaliações individualizadas, que classificam cada estudante, e afetam suas chances de entrar para uma universidade ou conseguir um bom trabalho. Neste sentido, elas são avaliações de *high stakes,* no linguajar dos especialistas ("de alto risco", mal traduzindo). Seu formato clássico é o do *Abitur* alemão, do *Bac* francês e de outros exames semelhantes existentes, sobretudo na Europa, como a *Maturità* italiana. Nos Estados Unidos, aonde a preocupação com os padrões e a qualidade do ensino médio é mais recente, foram desenvolvidas novas metodologias de avaliação baseadas em testes padronizados. A iniciativa partiu das universidades privadas, que sentiram a necessidade de um padrão externo de referência para a seleção de seus estudantes, que pudesse ser combinado com seus próprios critérios de escolha. Por sua iniciativa, foi criado o *College Entrance*

Examination Board, que, a partir de 1926, começou a desenvolver uma prova padronizada de múltipla escolha, o *Scholastic Aptitude Test* (SAT). Esta prova tem sido adotada pelo *College Board* desde 1941, e os resultados obtidos desde então são considerados como comparáveis através dos anos. Em 1947 foi criado o *Educational Testing Service*, que administra o SAT, e em 1959 foi criado um teste semelhante, o *American College Test*, requerido sobretudo pelas universidades públicas.[4]

Tanto o SAT como o ACT são exames que buscam avaliar a capacidade dos alunos de se saírem bem em seus cursos universitários, e não, necessariamente, se eles aprenderam os conteúdos específicos dados nos cursos de nível médio. São exames voluntários, mas requeridos pela maioria das instituições de nível superior, como parte de seus processos de seleção de alunos. Cerca de dois terços dos estudantes que terminam a escola secundária nos Estados Unidos fazem um destes testes, que são solicitados por cerca de três mil instituições.

Estes exames são consagrados, mas nem por isto deixam de receber críticas. Por exemplo:

> Los resultados del SAT han sido socialmente aceptados como el medio que permite elegir la institución de educación superior donde se ha de estudiar. Sin embargo, a partir de los años setenta, diversas organizaciones de derechos civiles y agrupaciones feministas han cuestionado el test porque contiene sesgos discriminatorios para grupos étnicos y mujeres. Por otra parte, las críticas a las preguntas van en aumento en los medios académicos: (i) los ítems se orientan a "reconocer" información y no miden la capacidad para "generarla", (ii) existe la paradoja de que la capacidad de escribir en inglés se mide sin tener que escribir absolutamente nada, (iii) al existir una sola respuesta válida para cada ítem no hay lugar para elaborar respuestas complejas que demuestren habilidades mayores que el simple reconocimiento de información, (iv) la brevedad del tiempo para resolver tantas preguntas concede más importancia a pensar rápido antes que a pensar profundamente, lo que hace que durante la preparación para la prueba se desarrollen "artimañas" para descartar las respuestas incorrectas, etcétera. (McLauchlan de Arregui et al. 2001), p. 34 (anexo 1).

[4] Estas informações, assim como as de outros sistemas de avaliação da conclusão do ensino médio, estão no anexo 1 de (McLauchlan de Arregui et al. 2001).

Apesar das críticas[5], os resultados do SAT, além de servirem para avaliar cada estudante individualmente, servem também de critério para a avaliação de cursos, instituições e do sistema educativo como um todo. Universidades de maior prestígio e reconhecimento exigem escores mais altos no SAT de seus estudantes do que outras, e estes resultados são públicos; escolas secundárias utilizam os resultados médios de seus alunos no SAT para saber se seu ensino é de alto ou baixo padrão; e os próprio governos podem usar os resultados médios da população, ou de grupos populacionais específicos, para estabelecer políticas de melhoria do desempenho escolar. Assim, se determinados grupos sociais obtêm resultados sistematicamente piores no SAT, como os negros nos Estados Unidos, uma possibilidade seria questionar a possível tendenciosidade do teste em favor da cultura dominante, da população branca; a outra possibilidade seria aceitar os resultados, e trabalhar para identificar as causas das diferenças e superá-las.

A metodologia norte-americana tem sido adotada por outros países, como o Chile e o Brasil. O Chile, em 1966, eliminou o antigo "bachillerato"[6] em moldes europeus, e a partir de então o Conselho Nacional de Reitores começou a implementar uma "Prueba de Aptidud Académica" (PAA), nos moldes do SAT, por iniciativa do Conselho de Reitores das Universidades Chilenas. Além de servir como critério para a seleção dos candidatos às universidades, as médias dos PAAs dos alunos admitidos são utilizadas pelo governo como critério para aferir a qualidade das universidades, e contam pontos para a distribuição de subsídios públicos para seu funcionamento. O Brasil, desde 1998, vem implementando o Exame Nacional de Ensino Médio (ENEM), como exame voluntário e por iniciativa do Instituto Nacional de Estudos e Pesquisas Educacionais (INEP), um órgão do Ministério da Educação. O objetivo é fazer dele um padrão nacional para a admissão de estudantes às universidades, e que possa também servir de critério para a

---

[5] Em 2001, depois de uma forte crítica do presidente da Universidade da California, que propôs que a Universidade abandonasse o SAT em benefício de testes de conhecimentos específicos, o College Board se comprometeu em mudar o teste, para torná-lo mais adequado. Veja a respeito (Gehring 2001; Beatty, Greenwood, and Linn 1999)

[6] O termo "bachillerato" tem sido usado na América Latina tanto no sentido de uma prova que é feita ano final do curso secundário, como no *Bac* francês, como no sentido de cursos que preparam alunos para estes cursos, como Gymnasium alemão.

avaliação das próprias escolas secundárias aonde os alunos se formam. Segundo o Ministério da Educação, em 2002,

> Os resultados da quinta edição do Enem consolidam a avaliação como um importante instrumento para a reforma do ensino médio no Brasil. O Exame contou com a participação de 1,3 milhão de jovens do total de 1,8 milhão que havia se inscrito. Desse total, 87% foram beneficiados com a isenção da taxa de inscrição. Até o momento, 384 instituições de ensino superior, das quais 44 públicas, utilizam o Enem nos seus processos seletivos aos cursos de graduação. (Brasil Ministério da Educação and Instituto Nacional de Estudos e Pesquisas Educacionais 2002).

Finalmente, os exames de conclusão do ensino médio funcionam como uma "identidade internacional de competência" cuja importância cresce à medida em que a circulação internacional de estudantes também aumenta. O SAT americano é oferecido em todo o mundo para quem se disponha a pagar, e seus resultados são reconhecidos por instituições americanas e de outros países. O *Abitur* e o *Bac* francês são reconhecidos internacionalmente, e existe um *International Baccalaureate* coordenado por uma fundação suíça que é dado por mais de mil instituições em todo o mundo – não só os exames, mas sobretudo os cursos que conduzem a ele. Os resultados do ENEM brasileiro ainda não têm reconhecimento claro, mas é provável que sua importância aumente cada vez mais, como indicador de desempenhos dos alunos que completam a educação secundária em qualquer parte do país. Assim, apesar das críticas que se possam fazer a estes sistemas de avaliação, a tendência é que eles continuem a existir e a se aperfeiçoar, como exigência de um mundo integrado onde a competência, e seu reconhecimento, jogam um papel cada vez mais central.

As avaliações da educação básica

Na educação básica, os estudantes são avaliados regularmente pelos seus professores, e, no final de cada ano letivo, correm o risco de serem aprovados ou não, e, no final do curso, receberem ou não seus diplomas. Ao lado destas avaliações tradicionais, existem agora os procedimentos sistemáticos e controlados de avaliação da educação básica em seu conjunto, para países e regiões, nos moldes do NAEP americano e do SAEB brasileiro, já mencionados. O principal objetivo destas avaliações não é a certificação dos alunos, como no caso das de nível secundário, mas das próprias instituições. Neste sentido, elas são avaliações de baixo risco (*low stakes*) para os alunos, mas podem ser de alto risco para redes escolares, escolas isoladas, seus

dirigentes e professores, já que conduzem a um veredicto a respeito de seu desempenho, que pode ter conseqüências importantes.

O ponto de partida destas avaliações é o "National Assessment of Educational Progress" americano, NAEP. Conforme sua descrição oficial,

> Since 1969 the National Assessment of Educational Progress (NAEP) has been regularly conducting assessments of samples of the nation's students attending public and private schools at the elementary, junior high, and high school levels. NAEP's goal, since its inception, has been to make available reliable information about the academic performance of American students in various learning areas. To this end, NAEP has produced more than 200 reports in 11 instructional areas. Teachers, administrators, and researchers from across America have helped propel NAEP into the valuable informational source it is today. As a result, members of the educational community are able to make use of NAEP's findings on students' learning experiences to inform policymakers and to improve students' educational experiences. NAEP is an indicator of what students know and can do. Only group statistics are reported, no individual student or teacher data are ever released.

O Sistema de Avaliação da Educação Básica do Brasil (SAEB) segue a metodologia do NAEP, e também produz resultados para cada Estado do país, sem, no entanto, proporcionar informações mais detalhadas por regiões, municípios, escolas e muito menos alunos. Ele permite dizer se os alunos de um Estado estão, na média, se saindo melhor do que os de outra parte, e, na medida em que os dados são comparáveis de ano a ano, permite dizer se cada estado, e o país como um todo, estão fazendo progresso. Além de uma informação técnica, este dado, ao ser divulgado, cria uma pressão política sobre as autoridades educacionais dos Estados, para que busquem resultados mais satisfatórios.

Além de ordenar os Estados ao longo de uma escala de desempenho, o SAEB permite relacionar determinados níveis de conhecimento com determinados conteúdos. Assim, o relatório do SAEB de 2001 mostra que "a média dos alunos brasileiros da 8ª série do ensino fundamental em Língua Portuguesa era 235, localizada no Nível 4 da escala de desempenho. Nesse nível, estão posicionados cerca de 36% dos alunos avaliados pelo Saeb 2001 nessa série e disciplina." E diz mais adiante:

Seria desejável que um maior percentual de alunos se concentrasse nos níveis seguintes da escala, especialmente no Nível 5, no qual se localizam cerca de 28% dos alunos, e no Nível 6, em que estão localizados aproximadamente 10% dos alunos. Cumpre destacar que apenas neste último nível os alunos adquirem a competência para operar com textos mais complexos, identificando o tema e os sentidos metafóricos em textos narrativos longos, como contos.

O SAEB permite ainda buscar correlações entre os resultados dos exames e as características das escolas e dos alunos, graças a informações colhidas através de questionários que acompanham os exames. A análise destas informações permite aprofundar o entendimento sobre fatores de ordem social ou institucional que afetam o desempenho dos estudantes, servindo de base, portanto, para políticas de melhoria de qualidade mais focalizada.

A extensão natural desta metodologia seria proporcionar informações sobre o desempenho de cada escola e mesmo de cada estudante, de tal forma que os problemas e as deficiências existentes pudessem ser localizadas com precisão. Alguns Estados brasileiros, como Paraná, São Paulo Bahia e Minas Gerais, desenvolveram seus próprios sistemas de avaliação, que proporcionam informações sobre cada escola, ou a maior parte delas. No Paraná, desde 2001, estas informações são consolidadas em um "Boletim da Escola", que é distribuído às autoridades escolares, professores e pais dos alunos de cada escola, com a esperança de que isto os motive a pressionar e trabalhar por melhores resultados. Na Bahia, deu-se início a um programa contínuo de avaliação do desempenho dos alunos ao longo do ano, que pudesse permitir intervenções em tempo real, para corrigir deficiências e problemas antes que eles se cristalizem. Nos Estados Unidos, a criação de standards nacionais de desempenho e de testes padronizados para os alunos que concluem a educação básica têm sido objeto de vários experimentos e intensa discussão.

As críticas e discussões que circundam o NAEP, o SAEB e suas derivações são semelhantes às que afetam ao SAT[7]. Provas quantitativas não

---

[7] A respeito, o Programa de Promoção da Reforma Educativa para a América Latina e o Caribe (PREAL) e o Grupo de Estudios del Desarrollo do Peru mantém um grupo de trabalho e discussão permanente sobre o tema dos estândares e avaliação da educação, que pode ser acessado pela Internet em http://www.grade.org.pe/gtee-preal/gteepreal.htm .

permitiriam medir os aspectos mais qualitativos da educação e da aprendizagem; fariam com que os estudantes se preparassem para a prova, e não para realmente aprender; as provas seriam tendenciosas, produzindo resultados sistematicamente piores para minorias culturais, lingüísticas ou raciais; elas sobrecarregariam as escolas com testes sucessivos, sem nenhum benefício pedagógico para os estudantes. Além destas questões técnicas, cujo mérito não caberia examinar aqui, estes exames trazem um forte componente político, na medida em que a avaliação das escolas é transferida das mãos dos professores e funcionários das secretarias de educação, em geral educadores e pedagogos, para especialistas em estatística e psicometria, gerando políticas gestadas fora das escolas, que podem afetar seus destinos.

Além de sua extensão e detalhamento para dentro dos países e regiões, estes exames estão sendo também aplicados a nível internacional, por iniciativas de agências como a Organização para a Cooperação Econômica e o Desenvolvimento (OECD) e a UNESCO (OECD Programme for International Student Assessment 2001). Estas comparações mostram profundas diferenças de resultados, muitas vezes surpreendentes e incompatíveis com o nível de renda e desenvolvimento dos países, permitindo, em princípio, mostrando que a qualidade da educação não depende somente de recursos, mas de características institucionais, pedagógicas e culturais dos países, que precisam ser melhor compreendidas.

A avaliação do ensino superior e da pós-graduação.

As universidades sempre se debateram entre os pólos da autonomia e do controle externo. Como instituições responsáveis pela criação e transmissão das formas mais complexas de conhecimento e cultura, organizadas e desenvolvidas na tradição das corporações medievais, as universidades gostam de se ver e ser consideradas como instituições autônomas, auto-reguladas, que não teriam como ser controladas e supervisionadas por pessoas e instituições que são, por definição, menos capazes e competentes do que elas mesmas. Por outro lado, desde seus primórdios, as universidades foram provedoras de serviços, e como tais deviam render contas a seus clientes, os estudantes, pela qualidade dos serviços que proporcionavam. Elas também desempenhavam funções públicas, ao formar e certificar os médicos, prelados

e juristas, cujo exercício profissional dependia da regulação e controle do Estado, da Igreja e da corporações profissionais.

Estes três grandes condicionantes definem os vértices do que ficou conhecido na literatura como o “triângulo de Clark”, o espaço de poder e autoridade criado pela corporação, pelo Estado e pelas demandas de mercado, dentro do qual as instituições universitárias se situam. A estes três pólos há que acrescentar um quarto, de grande importância na história da América Latina, que é a política (Clark 1983). Basta olharmos em volta para vermos como este espaço se constitui, e como as instituições tratam de se colocar da melhor maneira possível em seu interior: as universidades públicas, governadas pelos seus Conselhos, Departamentos e Congregações; as universidades privadas, buscando responder da melhor forma possível às demandas do mercado; o Ministério da Educação, tentando controlar os orçamentos, estabelecer critérios de qualidade, e regular a expansão; os partidos políticos e a sociedade, demandando por mais acesso, mais equidade, mais recursos públicos, e menos custos privados. Nenhuma instituição é imune a estas três influências, todas têm que responder, de alguma forma, às pressões e regulações dos governos, dos mercados, da política e de sua própria comunidade.

Quando o ensino superior se restringia dois ou três por cento da população, o vértice corporativo geralmente predominava, em diferentes formas de cooperação e entendimento com as burocracias governamentais. Com o ensino superior de massas, atendendo a proporções que vão de quinze por cento a mais da metade da juventude, os custos do ensino superior aumentam, as expectativas da sociedade e dos movimentos sociais sobre seus resultados se intensificam, os mecanismos tradicionais de coordenação e supervisão burocrática e corporativa já não funcionam, e as pressões por avaliações externas, transparentes e sistemáticas aumentam. Os estudantes querem saber quais são os bons cursos e as boas universidades, para poderem decidir melhor sobre seu futuro; os governos, pressionados por custos crescentes, querem saber porque um estudante de engenharia na universidade “a” custa três vezes mais do que na universidade “b”; os políticos gostariam de usar os recursos das universidades para outros projetos de sua preferência; os professores do ensino médio não entendem porque ganham tão menos do que os da universidade; empresários vislumbram grandes negócios na “indústria

do conhecimento", e se opõem às reservas de mercado e controles burocráticos que limitem suas atividades (Altbach 1999); conselhos de medicina, direito e engenharia se preocupam com a expansão desregulada de seus mercados de trabalho, e a proliferação de "fábricas de diplomas"; organizações não governamentais de minorias demandam quotas para seus públicos, e políticos das regiões mais pobres se queixam de que os recursos estão concentrados nas regiões mais ricas.

Além desta confluência de interesses e pressões contrapostos, as instituições de ensino superior têm funções mais gerais, de interesse da sociedade como um todo, sobre as quais os governos devem zelar: a formação de técnicos e especialistas bem qualificados, a pesquisa acadêmica e aplicada, a formação pós-secundária de massas, a educação continuada. Nem todas as instituições podem fazer igualmente todas estas coisas, todas elas custam dinheiro, e podem beneficiar a determinados grupos em relação a outros (Castro and Levy 2000)

A tudo isto se acrescenta uma nova dimensão, que é a internacionalização crescente da educação superior. Não se trata, a rigor, de uma novidade, porque as universidades são, desde os tempos medievais, instituições que transbordam as fronteiras dos países, quer no recrutamento de seus alunos, quer na contratação de seus professores, quer na influência que exercem, além de suas fronteiras nacionais. Mas, na entrada do terceiro milênio, a internacionalização se torna cada vez mais intensa. Estudantes, universidades, empresas e governos precisam saber se um diploma de médico ou de economista dado no Brasil vale o mesmo do que um outro dado em Portugal ou nos Estados Unidos; universidades e empresas estabelecem campus avançados em outros países, e gostariam de ter seus diplomas reconhecidos; cursos à distância e semi-presenciais, dados por instituições privadas ou públicas, não respeitam fronteiras, e disputam espaço com instituições nacionais.

Muitos mecanismos têm sido propostos para responder a esta pluralidade de demandas e pressões, alguns mais bem sucedidos que os outros. O mais conhecido, talvez, seja o sistema de avaliação e acompanhamento do desempenho das universidades inglesas, criado na época do governo Thatcher, e que ainda hoje vigora naquele país (Williams 1997). As universidades

passaram a ser submetidas a avaliações quantitativas permanentes, tanto de seus custos quanto de seu desempenho; foram forçadas a competir por recursos públicos e privados; e tiveram que alterar seus sistemas de gestão, aproximando-se cada vez mais às formas gerenciais do setor privado. Esta experiência foi marcada por um clima fortemente anti-acadêmico e contra o setor público que prevalecia na Inglaterra naqueles anos, e tem sido até hoje questionada pelo seu impacto negativo sobre a autonomia e a liberdade acadêmica. Ao mesmo tempo, permitiu, ou foi acompanhada, de uma forte expansão do acesso à educação superior na Inglaterra, e da eliminação do antigo sistema "binário" que separava as universidades das escolas politécnicas naquele país. Em sua avaliação, Gareth Williams mostra que, graças à reforma, o custo por estudante na Inglaterra caiu dramaticamente, sendo hoje 40% do que era vinte anos atrás em termos reais. Ao mesmo tempo, o acesso ao ensino superior cresceu, e a qualidade da pesquisa e do ensino parecem ter melhorado. Por outro lado, as atividades burocráticas dos professores ingleses ocupam hoje muito de seu tempo, e os salários não melhoram, criando sérios problemas de perda dos melhores talentos para o setor privado ou para as universidades americanas.

A experiência inglesa tem tido um impacto muito grande em todo o mundo, com várias tentativas mais ou menos bem sucedidas de cópia ou adaptação.[8] Os Estados Unidos têm uma tradição de avaliação bastante diferente. Sua base são um conjunto de agências de acreditação das universidades, estabelecidas por elas mesmas através de consórcios, que realizam visitas periódicas às instituições, e avalizam seu *status* acadêmico e sua qualidade.[9] Além disto, existe nos Estados Unidos toda uma indústria privada de avaliações "reputacionais", feita por empresas privadas através de questionários e entrevistas com diferentes setores. Estas avaliações classificam as instituições de nível superior segundo uma série de critérios, da qualidade

---

[8] O exemplo mais significativo na América Latina é o do Chile, com a reforma introduzida durante o regime Pinochet, que foi mantida com poucas alterações pelos governos da "concertación" que o sucederam. A reforma na Holanda parece ter sido mais bem sucedida do que na Inglaterra, graças à participação mais ativa da comunidade acadêmica em sua implementação. Veja, para o Chile, (Brunner 1997); para a Holanda, (Vught 1997).

[9] Sobre o sistema americano, (El-Khawas 1998). No início do século XX, o ensino médico americano, até então desregulado e caótico, passou por uma profunda avaliação, que teve como consequência implantar um forte padrão de qualidade, e restringir sua expansão. Veja a respeito (Flexner 1968); e (Schwartzman 1993).

da pesquisa à qualidade da cerveja, s são consumidas pelos milhões de estudantes que entram no ensino superior a cada ano.

No Brasil, durante muito tempo,a única avaliação disponível sobre os cursos superiores no Brasil era feita anualmente pela revista *Playboy*. A necessidade de um sistema moderno de avaliação sistemática do ensino superior tem sido levantada desde, pelo menos, os trabalhos da Comissão Nacional de Avaliação do Ensino Superior de 1985 (Brasil Ministério da Educação 1985), e deu origem a uma série de iniciativas desde então (Paul, Ribeiro, and Pillati 1992), até culminar nos procedimentos adotados pelo Ministério da Educação na gestão de Paulo Renato de Souza, em duas partes principais. A primeira, bastante tradicional, é a chamada "avaliação das condições de oferta", em que comissões de professores visitam os estabelecimentos e emitem juízo sobre os recursos humanos e físicos de que os cursos dispõem; a segunda, bastante pioneira, é o Exame Nacional de Cursos, conhecido como "Provão", que consiste em um exame padronizado aplicado aos alunos que concluem os cursos superiores em todo o país em diferentes áreas de conhecimento.

Apesar de ser uma prova individual, o provão não é utilizado para certificar o aluno, e sim para avaliar os cursos, pela média do desempenho de seus formados. Este sistema tem alguns problemas conhecidos, como o de impor um padrão único a todas as instituições; o de informar somente a distribuição estatística dos conceitos, sem definir padrões mínimos de qualidade; e o de não medir o "valor agregado" dos cursos, e sim os resultados absolutos, que podem estar fortemente influenciados pelas condições prévias dos alunos, favorecendo assim as instituições de elite. Além disto, não existem praticamente conseqüências para os cursos que são mal avaliados. Apesar disto, o "provão" gerou um clima de discussão e preocupação com a qualidade da educação superior que os antigos procedimentos burocráticos e administrativos de avaliação e credenciamento de cursos não conseguiam produzir, e já começam a influenciar as escolhas dos estudantes.

Finalmente, o Brasil possui também um sistema bastante original e bem estabelecido de avaliação de seus cursos de pós-graduação. Na Europa e nos Estados Unidos, a pós-graduação é avaliada seja como parte das avaliações das instituições como um todo, seja por pesquisas reputacionais, ou

pela avaliação do desempenho científico e acadêmico de seus professares e pesquisadores. No Brasil, a pós-graduação foi implantada muito rapidamente, a partir da reforma universitária de 1968, e os mecanismos tradicionais e autorização de cursos através do Conselho Federal de Educação não conseguiam acompanhar seu crescimento. O sistema estabelecido a partir da década de 70, e em vigência até hoje, consistiu na criação de comitês de especialistas que fazem avaliações qualitativas de cada curso ou programa, trabalhando a partir de indicadores quantitativos como número de alunos formados, qualificação dos docentes e qualidade e quantidade da pesquisa científica e tecnológica produzidas (Castro and Soares 1986).

Estas avaliações se tornaram necessárias pela proliferação dos cursos de pós-graduação no país, que ocorreu como resposta às exigências de credenciais acadêmicas criadas pelo governo federal para os professores universitários. Ao contrário do que ocorre com as avaliações da educação superior, as avaliações da CAPES têm conseqüências importantes para os cursos, não só pelo prestígio associado a seus conceitos, mas também pelos recursos que são distribuídos em função destes conceitos.

Conclusão: o presente e o futuro das avaliações

As avaliações de nova geração que surgiram em todo o mundo nas últimas décadas são tentativas mais ou menos bem sucedidas de responder a problemas reais, e deverão continuar a existir, apesar de seus eventuais problemas e dificuldades. Não é que as avaliações consigam atender a todas as questões que preocupam; mas os problemas associados às avaliações parecem ser claramente preferíveis aos problemas associados à sua não existência, como a falta de parâmetros, a ausência de critérios, os desperdícios e injustiças na distribuição de recursos, e a impossibilidade de estabelecer políticas consistentes para sistemas de educação de massas. Dos problemas existentes, muitos são de natureza técnica, e podem ir sendo superados na medida em que os estudos, as pesquisas e as metodologias de avaliação se desenvolvam. Assim, por exemplo, a crítica costumeira de que as medidas “quantitativas” não teriam como avaliar aspectos qualitativos da educação, como a capacidade de leitura e escrita, pode ser superada com instrumentos bem feitos e calibrados; e já existem sistemas computadorizados capazes de avaliar textos

em larga escala e de forma mais consistente, confiável do que o que fariam avaliadores e carne e osso.

Mas a mensuração do desempenho é só uma das funções da avaliação educacional. Desde o Coleman Report, as avaliações servem também como instrumento para identificar os fatores que afetam o desempenho escolar, e para avaliar o resultado ou fracasso das políticas educacionais. Por exemplo, as avaliações têm permitido medir, com precisão, o peso relativo das condições de vida das famílias dos estudantes, do funcionamento das escolas e do trabalho dos professores nos resultados obtidos. Estas análises tem mostrado, por exemplo, como é importante que a família se envolva na vida escolar dos estudantes; o papel negativo da burocratização da administração escolar; e os efeitos pedagógicos das diferentes orientações e níveis de formação dos professores.

As grandes controvérsias associadas a estas avaliações têm menos a ver com suas dificuldades técnicas do que com suas conseqüências e implicações políticas. As avaliações de nova geração reduzem o poder e a autoridade dos educadores, colocando em seu lugar uma nova geração de especialistas com outro tipo de cultura e informação. Os resultados das avaliações podem afetar situações estabelecidas de poder e estabilidade institucional, ameaçando carreiras, e transferindo recursos de determinadas instituições para outras. Elas criam, necessariamente, um ambiente competitivo entre pessoas e instituições, que podem gerar ansiedade. E podem também, com facilidade, criar um novo tipo de burocracia capaz de gerar grandes custos e sofrimento para as instituições educacionais e seus participantes, sem necessariamente produzir os resultados benéficos que os defensores das avaliações pretendem ou apregoam.

Não há soluções fáceis para estes problemas, mas é bastante claro que as avaliações, e os avaliadores, não podem se eligir em donos absolutos de uma nova verdade, em substituição à autoridade dos pedagogos e acadêmicos do passado. As avaliações precisam ser constantemente avaliadas, e para isto alguns ingredientes são essenciais. Suas metodologias precisam se tornar explícitas, sem se esconder por traz do jargão especializado de estatísticos e psicometristas. A área acadêmica de estudo e pesquisas sobre os temas e metodologias de avaliação precisa ser ampliada e fortalecida, abrindo espaço

para perspectivas distintas e concorrentes. Os dados produzidos pelas avaliações devem ser de domínio público, e não ficar retidos como recursos de poder nas agências avaliadoras. Tanto quanto possível, as avaliações não devem ser únicas, mas produzidas por diferentes agências, com critérios e metodologias próprias, respondendo distintos setores – estudantes, associações profissionais, governos, conselhos de reitores, federações de escolas; e estas agências devem prestar contas, de forma responsável, às diferentes comunidades a que devem atender.

Referências

Altbach, Philip G. 1999. *Private Prometheus private higher education and development in the 21st century*, *Contributions to the Study of Education*. Westport, Conn: Greenwood Press.

Beatty, Alexandra, M.R.C. Greenwood, and Robert L. Linn, eds. 1999. *Myths and tradeoffs - the role of tests in undergraduate admissions*. Edited by National Academy of Sciences, Steering Committee for the Workshop on Higher Education Admissions, Board on Testing and Assessment, Commission on Behavioral and Social Sciences and Education and Office of Scientific and Engineering Personnel. Washington: National Academy Press. http://www.nap.edu/html/myths_tradeoffs/

Bourdieu, Pierre, and Jean Claude Passeron. 1970. *La reproduction; éléments pour une théorie du système d'enseignement*. [Paris]: éditions de Minuit.

Brasil Ministério da Educação. 1985. *Uma Nova Política para a Educação Superior Brasileira - Comissão Nacional para a Reformulação da Educação Superior - Relatório Final*. Simon Schwartzman (relator) ed. Brasilia: Imprensa universitária da Universidade Federal do Ceará.

Brasil Ministério da Educação, and Instituto Nacional de Estudos e Pesquisas Educacionais. 2002. *Mais de 70% dos participantes do ENEM têm desempenho de "regular a bom" na redação* 2002 [cited 19 de dezembro 2002]. Available from http://www.inep.gov.br/imprensa/noticias/enem/news02_16.htm.

Brunner, José Joaquín. 1997. From state to market coordination the Chilean case. *Higher Education Policy* 10 (3/4):225-237.

Castro, Cláudio de Moura , and Gláucio A. D. Soares. 1986. As avaliações da Capes. In *Pesquisa universitária em questão*, edited by S. Schwartzman and C. d. M. Castro. São Paulo: Editora da UNICAMP.

Castro, Cláudio de Moura, and Daniel C. Levy. 2000. *Myth, reality, and reform higher education policy in Latin America*. Washington: Inter-American Development Bank. Distributed by Johns Hopkins University Press.

Clark, Burton R. 1983. *The higher education system academic organization in cross-national perspective*. Berkeley: University of California Press.

Coleman, James Samuel, Ernest Q. Campbell, Carl F. Robson, James McPartland, and Alexander McFarlane Mood. 1966. *Equality of educational opportunity*. Washington: Department of Education.

Dasgupta, Partha, and Ismail Serageldin, eds. 2000. *Social capital a multifaceted perspective*. Washington, D.C: World Bank.

Durkheim, Emile. 1967. *De la division du travail social*. 8e ed, *Bibliotheque de philosophie contemporaine*. Paris,: Presses universitaires de France.

Durkheim, Emile, and Maurice Halbwachs. 1938. *L'evolution pedagogique en France*, *Bibliotheque de philosophie contemporaine*. Paris,: F. Alcan.

El-Khawas, Elaine. 1998. Accreditation's Role in Quality Assurance in the United States. *Higher Education Management* 10 (3):43-56.

Flexner, Abraham. 1968. *Universities American, English, German*. New York: Oxford University Press.

Fukuyama, Francis, and IMF Institute. 2000. *Social capital and civil society*, *IMF working paper*. Washington, D.C.: International Monetary Fund, IMF Institute.

Gehring, John. 2001. UC President pitches plan to end use of SAT in admissions. *The Chronicle of Higher Education*, February 28, 5.

Herrnstein, Richard J., and Charles A. Murray. 1994. *The bell curve: intelligence and class structure in American life*. New York: Free Press.

McLauchlan de Arregui, Patricia, Ernesto Melgar, Iván Montes, Simon Schwartzman, and Juan Fidel Zorrilla. 2001. *Certificación de aprendizajes escolares al término de la secundaria - Experiencias internacionales y opciones para el Perú*, *Programa Especial de Mejoramiento de la Calidad de la Educación Peruana - Documentos de Trabalho 4*. Lima: Ministerio de Educación del Perú. http://www.schwartzman.org.br/simon/grade3.pdf

National Assessment Governing Board. 2002. *About NAEP* 2002 [cited December 19 2002]. Available from http://www.nagb.org.

OECD Programme for International Student Assessment. 2001. Knowledge and skills for life - first results from PISA 2000 - Education and skills. Paris: OECD.

Parsons, Talcott. 1991. *The social system*. New ed ed, *Routledge sociology classics*. London: Routledge.

Parsons, Talcott, and Robert Freed Bales. 1966. *Family, socialization and interaction process*. New York: Free Press.

Paul, Jean-Jacques, Zoya Ribeiro, and Orlando Pillati. 1992. As iniciativas e as experiências de avaliação do ensino superior - balanço crítico. In *Avaliação do ensino superior*, edited by E. R. Durham and S. Schwartzman. São Paulo, Brasil: EDUSP.

Putnam, Robert D, Robert Leonardi, and Raffaella Nanetti. 1993. *Making democracy work civic traditions in modern Italy*. Princeton, N.J.: Princeton University Press.

Schwartzman, Simon. 1993. A qualidade no espaço universitário: conceitos, modelos e situação atual. Paper read at Congresso Internacional de Qualidade e Excelência na Educação, 25 a 29 de outubro, at Rio de Janeiro, Universidade Gama Filho. http://www.schwartzman.org.br/simon/gamaf.htm

The World Bank Group. 2002. *Public examination systems* 2002 [cited December 18 2002]. Available from http://www1.worldbank.org/education/exams/nature/default.asp#A%20Brief%20History.

United States Department of Education, and The National Commission on Excellence in Education. *A Nation at risk: the imperative for educational reform. A report to the Nation and the Secretary of Education* 1983 [cited. Available from http://www.ed.gov/pubs/NatAtRisk/index.html.

Vught, Frans A. van. 1997. Combining planning and the market: an analysis of Government strategy towards higher education in the Netherlands. *Higher Education Policy* 10 (3/4):211-224.

Williams, Gareth. 1997. The market route to mass higher education: British experience 1979-1996. *Higher Education Policy* 10 (3/4):275-290.

Wolf, Alison. 2002. *Does education matter? myths about education and economic growth*. London: Penguin.